ISSN 1745-8587

School of Economics, Mathematics and Statistics

BWPEF 0605

Auto-Dependence Structure of Arch-Models: Tail Dependence Coefficients

Raymond Brummelhuis

May 2006

Birkbeck Working Papers in Economics & Finance

AUTO-DEPENDENCE STRUCTURE OF ARCH-MODELS: TAIL DEPENDENCE COEFFICIENTS

RAYMOND BRUMMELHUIS

Abstract. We study autodependence in ARCH-models by computing the auto-lower tail dependence coefficients and certain generalizations thereof, for both stationary and non-stationary time series. This study is inspired by financial risk-management issues, and our results are relevant for estimating probabilities of consecutive value-at-risk violations.

1. Introduction

One of the striking aspects of empirical financial return-series is volatility-clustering: although successive returns are roughly uncorrelated, in the well-known phrase of Mandelbrot [19], 'large changes tend to be followed by large changes - of either sign - and small changes tend to be followed by small changes'. Auto-dependence in time series is mostly studied in terms of auto-correlation, and for financial returns the squared returns typically exhibit an auto-correlation pattern which is reminescent of an ARMA-process. A popular class of financial econometrical models that capture this kind of behavior is that of the GARCH-models introduced by Engle [13] and Bollerslev [7]; see also [8], [14].

Although auto-correlation of squared returns provides a satisfactory qualitative explanation for volatility clustering, and for Mandelbrot's observation cited above, it does not always easily provide answers to questions related to the multi-variate distribution functions associated to the process, like for example the behavior of conditional quantiles at lags greater than 1 for a GARCH(1, 1) - cf. [9]. Motivated by recent developments in risk-management (see Embrechts, McNeil and Straumann [12],), we propose to study auto-dependence in time-series from the point of view of alternative, copula-based, dependence measures like rank-correlations, concordance measures or, in this paper, tail dependence coefficients. Indeed, as in risk-management, where linear correlation is the natural dependence measure for multi-variate normal, or more generally elliptic, linear models, but is less suitable for non-linear models or for more general classes of multi-variate distributions, we argue that autocorrelations, while adequate for linear processes, should, for non-linear processes like the GARCH, be supplemented by other measures of auto-dependence. As a further motivation to consider alternative auto-dependence measures in time series, linear autocorrelations are not always defined: in the case of a stationary GARCH only for a limited parameter range, and not at all if the i.i.d. innovations which drive the process do not posses a finite variance. Similarly, if one is interested in the squared process they should at least have a finite fourth moment. (This does not preclude studying sample autocorrelation functions of such processes; cf. [4] and its references).

As already mentioned, the particular dependence measure we will study in this paper is the lower tail dependence coefficient, and certain generalizations thereoff. The *lower tail dependence coefficient* $\lambda_{X|Y}$ of two random variables X and Y is defined as the limit, for $\alpha \to 0$, of the conditional probability that X is smaller than the α-th quantile of X, given that Y will be below its own α-th quantile; cf. [12]. As we will see, this measure will pick up non-trivial auto-dependence in a stationary ARCH, but, somewhat surprisingly, not in a non-stationary one. For this reason we will introduce certain generalizations, which we will call *generalized lower tail dependence coefficients*, a typical example of which would be the limit, for α tending to 0, of the probability that X will be smaller than its $\sqrt{\alpha}$-th quantile given that Y than its α-th quantile. Here, the square root can be replaced by more general functions of α tending to 0 at a suitably slower rate than α itself.

The reasons for considering these particular dependence measures instead of others like Kendall's τ or Spearman's ρ are, firstly, that since these are asymptotic quantities, they are more amenable to a complete theoretical analysis: in terms of copulas, to determine $\lambda_{X|Y}$ and its generalizations, we only need to understand the behavior of the copula of X and Y near the lower left corner point of its domain of definition, while τ and ρ require a global knowledge of the copula. Secondly, and from the point of view of applications more importantly, the lower tail dependence coefficient has a direct relevance for financial risk-management, since it can be interpreted in terms of value-at-risk: if $X = (X_n)_n$ is a stationary time-series, representing the daily returns of an investment portfolio, the (unconditional) daily value-at-risk at a confidence level of $1 - \alpha$ is simply the (absolute value of) the lower α-th quantile of X_n. The financial interpretation is that with a probability of $1 - \alpha$, daily (percentage) losses will be less than this value-at-risk. Under the Basle rules, for a financial institution to suffer losses on its market portfolio exceeding the value-at-risk at some specified confidence level has regulatory consequences; cf [3]. The lower tail dependence coefficient $\lambda_{X_{n+1}|X_n}$ will provide information on the probability of violating one's value-at-risk limit on two consecutive days, and similarly for other time lags.

Value-at-risk, as a risk-measure, has been criticized for two, closely related, reasons: it does not give an estimate of the size of the loss when it occurs, and, when considering more than one risky asset or portfolio of risky assets, it can fail to be sub-additive. In the terminology of [2], it is not coherent. (It is coherent when restricted to portfolios made up of jointly elliptically distributed assets, cf.[12]). The failure of subadditivity is less relevant if $(X_n)_n$ models a financial institutions entire market portfolio. We note in this respect that Berkowitz and O'Brien [5] report that a simple univariate ARMA + GARCH-model of the total portfolio return is at least as effective, if not more, for value-at-risk forecasting than the detailed structural value-at-risk models commonly used by banks. As regards the loss-size, this can be estimated by the expected shortfall, which is basically the expectation of X_n given that it is smaller than its α-th quantile. If properly defined for non-continuous distributions, expected shortfall can be shown to be coherent, cf. [1], [21]. We won't however consider expected shortfall here, but limit attention to quantiles and value-at-risk, this being at the moment the industry standard.

To investigate the relevance of lower tail dependence as auto-dependence measure for non-linear time-series, we do in this paper a detailed study of the particular, but representative, example of an ARCH(1)-model, the simplest of the GARCH-models. Our results are expected to extend to more general GARCH-processes, and in particular to GARCH(1, 1)'s. If $(X_n)_{n\geq 0}$ is an ARCH(1), we will compute the lower tail dependence coefficient $\lambda_{X_{n+k}|X_n}$ both for stationary processes (taking for X_0 a stationary solution of the ARCH(1)-equation), and for non-stationary ones having an a.s. initial value $X_0 = x_0 \in \mathbb{R}$. The latter are of interest for conditional value-at-risk estimation given today's return; cf. [15]. Under the assumption that ϵ_n is symmetric we will give, in theorem 2.2 below, an explicit expression for $\lambda_{X_n|X_{n+k}}$ in the form of a multiple integral involving the pdf f_ϵ of ϵ_n, the auto-regressive ARCH(1)-parameter a_1 and the Pareto exponent of the stationary distribution (which can be computed from the former two), but no other parameters. In particular, more detailed knowledge of the stationary distribution itself is not required.

The non-stationary case is more delicate, and requires additional hypotheses on ϵ_n. We will suppose that f_ϵ, and its first two derivatives have tails with a Pareto-type inverse power decay. This class of f_ϵ was singled out because of its importance in empirical modelling: see e.g. [15]. Somewhat surprisingly, the tail dependence coefficients $\lambda_{X_{n+k}|X_n}$ now turn out to be 0, and to still be able to detect a non-trivial auto-dependence in the tails, we have to use the generalized coefficient mentioned above (see definition 2.5 below). The generalized lower tail dependence coefficients turn out to be all equal to 1/2, both for stationary and non-stationary processes. Their value is in particular independent of the ARCH(1)-parameters, which contrasts with the traditional picture provided by the linear auto-correlations of the squared process.

The paper is organized as follows: in section 2 we recall a number of basic definitions and facts, define the generalized tail dependence coefficients and state our main results, which are theorems 2.2 and 2.7 for stationary ARCH(1)'s, and theorems 2.4 and 2.6 for non-stationary ones. In section 3 we prove the results for stationary ARCH(1)'s. Sections 4 to 6 are devoted to the technically more complicated non-stationary case. We first determine, in section 4, the asymptotic behavior of the probability density function of X_n given an a.s. initial value for X_0. The main technical tool in this section is the Mellin transform. The asymptotics of section 4 are first used in section 5 to derive the asymptotics of the quantile function, and then, in section 6, the asymptotic behavior of the tail dependence function. Using these, our results for non-stationary ARCH(1)'s quickly follow. Finally, in section 7 we use Monte Carlo simulations to study lower tail dependence at small but non-zero values of the confidence parameter α. An appendix summarizes, for convenience of the reader, the relationship between Mellin transform and asymptotic expansions needed in sections 4 and 6.

2. Main Results

All random variables will be defined on some common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Recall that the left-inverse $F^{\leftarrow}$ of an increasing function $F : \mathbb{R} \to \mathbb{R}$ is defined by $F^{\leftarrow}(\alpha) = \inf\{x : F_X(x) \geq \alpha\}$ for $\alpha \in \mathbb{R}$;, cf. [6], [11]. If $F_X(x) = \mathbb{P}(X \leq x)$, $x \in \mathbb{R}$ is the cumulative distribution function of a random variable X, then the α-th quantile $q_X(\alpha)$ can defined as as the left-inverse of F_X; explicitly,

$$q_X(\alpha) = F_X^{\leftarrow}(\alpha) = \inf\{x \in \mathbb{R} : \mathbb{P}(X \leq x) \geq \alpha\}, \quad \alpha \in [0, 1].$$

All probability distributions we will encounter in this paper, or the functions by which we will be approximate them, will be continuous and strictly increasing in the regions of interest, so that the left-inverse reduces locally to the ordinary one, at least locally.

Let $(X_n)_{n\geq 0}$ be an ARCH(1)-process, defined by

$$X_{n+1} = \sqrt{a_0 + a_1 X_n^2}\, \epsilon_n, \tag{1}$$

where $a_0, a_1 > 0$, and $(\epsilon_n)_{n\geq 1}$ is an i.i.d. sequence of random variables with mean 0. We will mostly suppose, for symplicity, that ϵ_n is symmetric. We will consider both stationary and non-stationary ARCH(1)'s. The stationary case amounts to taking $X_0 =_d X_\infty$, where X_∞ is the stationary solution of (1). Stationary solutions exists if $a_0 > 0$ and $\mathbb{E}\big(\log a_1 \epsilon_1^2\big) < 0$, cf. [20] (reproduced in [14]), [17] or [11], chapter 8.4. Moreover, under mild additional hypothesis on ϵ_n, the stationary solution will have a Pareto-tailed probability distribution, cf. [18], [17], [16], and [11] for a text-book treatment of the case of normally distributed ϵ_n. Suppose that $\epsilon =_d \epsilon_n$ has a positive density such that for some $h_0 \in (0, \infty]$, $\mathbb{E}(|\epsilon|^h) < \infty$ for all $h < h_0$ and $\mathbb{E}(|\epsilon|^{h_0}) = \infty$, and let $\kappa_\infty = 2\mathrm{k}_\infty$, where $\mathrm{k}_\infty > 0$ is the unique positive solution to

$$\mathbb{E}\left((a_1\epsilon^2)^{\mathrm{k}_\infty}\right) = 1. \tag{2}$$

Then the probability distribution of the stationary solution X_∞ has a Pareto-type inverse power decay of the tails: there exists a constant $c_\infty > 0$ such that

$$F_{X_\infty}(x) \simeq \frac{c_\infty}{|x|^{\kappa_\infty}}, \quad x \to -\infty, \tag{3}$$

and similarly for the right tail $\overline{F}_{X_\infty}(x) = 1 - F_{X_\infty}(x)$. This result was recently generalized to arbitrary GARCH(p, q) in [4].

We next turn to tail dependence.

Definition 2.1. (cf. [12]) Let (X, Y) be a random vector. The *(lower) tail dependence function* of X on Y is the function $\lambda_{X|Y} : [0, 1] \to [0, 1]$ defined by

$$\lambda_{X|Y}(\alpha) := \mathbb{P}(X \leq q_X(\alpha) | Y \leq q_Y(\alpha)), \quad \alpha \in [0, 1]. \tag{4}$$

The *coefficient of lower tail dependence of X on Y* is defined by

$$\lambda_{X|Y} = \lim_{\alpha \to 0} \lambda_{X|Y}(\alpha), \tag{5}$$

provided the limit exists.

If the limit (5) does not exist, one can consider instead the lim sup and the lim inf, and interpret these as an upper, respectively lower bound on the dependence of X on Y in the extreme lower tail. If X and Y are independent, then $\lambda_{X|Y}(\alpha) = \alpha$ and $\lambda_{X|Y} = 0$, and if X and Y are *co-monotonic* or *counter-monotonic* (meaning that one is an increasing respectively decreasing function of the other), $\lambda_{X|Y}(\alpha) = \lambda_{X|Y} = 1$ respectively -1.

In the case of continuously distributed X and Y, $\lambda_{X|Y}(\alpha)$ will only depend on the *copula* $C_{X,Y}$ of X and Y, which in this case can be characterized as the unique function $C_{X,Y} : [0, 1] \times [0, 1] \to [0, 1]$ such that the joint probability distribution $F_{X,Y}$ can be written as

$$F_{X,Y}(x, y) = C_{X,Y}\left(F_X(x), F_Y(y)\right).$$

Since for continuous increasing F, $F \circ F^{\leftarrow} = id$, it follows that $\lambda_{X|Y}(\alpha) = \alpha^{-1} C_{X,Y}(\alpha, \alpha)$. The lower tail dependence coefficient $\lambda_{X|Y}$ can therefore be interpreted as the directional derivative, in the direction of the diagonal, of $C_{X,Y}$ in $(0, 0)$.

In financial modelling terms, thinking of X and Y as percentage returns of financial assets over some fixed common time-period, $-q_X(\alpha)$ is interpreted as the value-at-risk, VaR_α, with confidence $1 - \alpha$ over the time period under consideration (to be multiplied by the amount of capital invested, to be precise). In practice, α will be small, $\alpha = 0.05$ or 0.01, and $q_\alpha(X)$ will be negative; by convention, losses are recorded as non-negative numbers, whence the minus-sign. With this terminology $\lambda_{X|Y}(\alpha)$ is the probability of the losses of X exceeding $\mathrm{VaR}_\alpha(X)$, given that the losses of Y already exceed $\mathrm{VaR}_\alpha(Y)$. We refer to [12] for a detailed discussion of copulas and dependence measures, and their applications to risk management.

Our first result gives an expression for the autotaildependence coefficients of a stationary ARCH(1).

Theorem 2.2. *Let $(X_n)_{n\geq 0}$ be a strictly stationary ARCH(1) having symmetrically distributed innovations $\epsilon_n =_d \epsilon$ with density f_ϵ such that the stationary distribution has a Pareto-type tail decay (3). Let κ_∞ be defined by (2) and let F_ϵ be the cumulative distribution function of $\epsilon =_d \epsilon_n$. Then the coefficient of lower tail dependence of X_{n+p} on X_n is given by*

$$\lambda_{X_{n+p}|X_n} = \kappa_\infty \int_{-\infty}^{-1} \int_{\mathbb{R}^{p-1}} F_\epsilon \left(-\frac{1}{a_1^{p/2}|x_1|\cdots|x_{p-1}||z|} \right) \prod_{j=1}^{p-1} f_\epsilon(x_j)\ |z|^{-\kappa_\infty - 1}\, d\mathbf{x}dz,$$

where $d\mathbf{x} = dx_1 \cdots dx_{p-1}$. In particular, $\lambda_{X_{n+p}|X_n} \neq 0$.

The integral on the right is convergent, since F_ϵ is bounded by 1 and $\kappa_\infty > 0$. It is easily verified that its value is between 0 and 1. Note that $\lambda_{X_{n+p}|X_n}$ only depends on F_ϵ, a_1 and κ_∞. In particular, it does not depend on any further characteristics of the, in general unknown, stationary distribution F_{X_∞}, not even on the constant c_∞ in (3).

We next turn to non-stationary ARCH(1)'s starting with an a.s. initial condition $X_0 = x_0 \in \mathbb{R}$. Let $\mathbb{P}^{x_0}$ be the conditional probability $\mathbb{P}^{x_0} = \mathbb{P}(\cdot|X_0 = x_0)$ and let $F_X^{x_0}$, $q_X^{x_0}(\alpha)$ and $\lambda_{X|Y}^{x_0}$ denote, respectively, the cumulative distribution function, quantile and lower tail dependence coefficient with respect to $\mathbb{P}^{x_0}$. In the non-stationary case we will suppose that $\epsilon =_d \epsilon_n$ has a twice differentiable symmetric probability density f_ϵ whose derivatives up to order 2 satisfy a Pareto-type decay condition:

Condition 2.3. There exists constants $\kappa_\epsilon, c_\epsilon > 0$ such that

$$f_\epsilon(x) = \frac{c_\epsilon}{|x|^{\kappa_\epsilon+1}} + \rho_\epsilon(x), \quad x \neq 0, \tag{6}$$

with remainder ρ_ϵ satisfying

$$|\rho_\epsilon|,\ |x\frac{d}{dx}\rho_\epsilon|, |x^2\frac{d^2}{dx^2}\rho_\epsilon|\ \leq \frac{C}{|x|^{\kappa_\epsilon+1+\eta_0}}, \quad |x| > 0, \tag{7}$$

for suitable constants $C, \eta_0 > 0$.

We use $\kappa_\epsilon + 1$ as Pareto exponent in (6) since this corresponds to an exponent of κ_ϵ for F_ϵ. The hypotheses on f_ϵ cover cases like the Student distributions and the Pareto-type distributions originating from extreme value theory; cf. [12], [15]. It follows from (6) and (7) that the j-th derivative satisfies $f_\epsilon^{(j)}(x) = c_j|x|^{-\kappa_\epsilon - j - 1} + O\left(|x|^{-\kappa_\epsilon - j - 1 - \eta_0}\right)$, for $j = 0, 1, 2$. We then have:

Theorem 2.4. *Suppose the pdf f_ϵ of ϵ_n is symmetric and satisfies condition 2.3. Then the conditional tail dependence coefficients all vanish: $\lambda_{X_{n+k}|X_n}^{x_0} = 0$ for all $x_0 \in \mathbb{R}$, $n, k \geq 1$.*

This is a somewhat unexpected result, whose apparent contradiction with theorem 2.2 as $n \to \infty$ will be clarified in sections 5 and 6 below, cf. remark 6.2. There does not seem to be anything special about the ϵ_n's satisfying condition 2.3: using the results of [9], theorem 2.4 can be shown to be also true for GARCH(1, 1)'s with normal innovations

The message of theorem 2.4 is that from the point of view of the lower tail dependence coefficient there is no difference between a non-stationary ARCH(1) and a strict white noise process. This is counter-intuitive, and to still be able to quantify the difference between these two cases, we introduce the following generalisation of the lower tail dependence coefficient.

Definition 2.5. let (X, Y) be a real random vector, and let $\psi : (0, 1] \to (0, 1]$, such that $\lim_{\alpha \to 0} \psi(\alpha) = 0$. We define the *generalized lower tail dependence coefficient of* X *on* Y, $\lambda^{\psi}_{X|Y}$, by

$$\lambda^{\psi}_{X|Y} := \lim_{\alpha \to 0} \mathbb{P}\left(X \leq q_X(\psi(\alpha)) | Y \leq q_Y(\alpha)\right), \tag{8}$$

assuming the limit exists.

Only the behavior of ψ at 0 matters and ψ only needs to be defined on some small sub-interval $(0, \delta)$, $\delta > 0$. It is easily seen that for continuous F_X and F_Y, $\lambda^{\psi}_{X|Y}$ again only depends on the copula $C_{X,Y}$, and that

$$\lambda^{\psi}_{X|Y} = \lim_{\alpha \to 0} \frac{1}{\alpha} C_{X,Y}(\psi(\alpha), \alpha),$$

the directional derivative of $C_{X,Y}$ in $(0, 0)$ along the curve $\alpha \to (\psi(\alpha), \alpha)$. We specifically allow curves which are tangential to one of the axes of the copula's domain of definition, $[0, 1] \times [0, 1]$. If X and Y are independent, then clearly $\mathbb{P}\left(X < q_X(\psi(\alpha)) | Y < q_Y(\alpha)\right) = \psi(\alpha)$, and $\lambda^{\psi}_{X|Y} = 0$. Embrechts, McNeil and Strauman [12] call two random variables X and Y asymptotically independent in the lower tail if $\lambda_{X|Y} = 0$. More generally, one might call X and Y *strongly asymptotically independent* (in the lower tail) if $\lambda^{\psi}_{X|Y} = 0$, for all positive ψ on $(0, 1]$ for which $\lim_{\alpha \to 0} \psi(\alpha) = 0$. Theorem 2.4 showed that consecutive values X_n, X_{n+p} of a conditional ARCH(1) with a.s. initial condition are asymptotically independent in the lower-tail. The next theorem shows that they are far from being strongly asymptotically independent. Let $\lambda^{x_0,\psi}_{X|Y}$ be the twisted lower tail dependence coefficient (8) computed with respect to the measure $\mathbb{P}^{x_0}$.

Theorem 2.6. *Under the hypotheses of theorem 2.4, we have that*

$$\lambda^{x_0,\psi}_{X_{n+p}|X_n} = \frac{1}{2}, \tag{9}$$

for all functions $\psi : (0, 1] \to (0, 1]$ *satisfying*

$$\lim_{\alpha \to 0} \psi(\alpha) = \lim_{\alpha \to 0} \frac{\alpha (\log \psi(\alpha)^{-1})^{p-1}}{\psi(\alpha)} = 0. \tag{10}$$

There is an analogue of theorem 2.6 for stationary ARCH(1)'s:

Theorem 2.7. *Under the conditions of theorem 2.2, if* $(X_n)_{n \geq 0}$ *is a stationary ARCH(1), then* $\lambda^{\psi}_{X_{n+p}|X_n} = \frac{1}{2}$, *for all* $\psi = \psi(\alpha)$ *satisfying*

$$\lim_{\alpha \to 0} \psi(\alpha) = \lim_{\alpha \to 0} \frac{\alpha}{\psi(\alpha)} = 0.$$

Typical examples of functions $\psi(\alpha)$ satisfying the conditions of theorems 2.6 and 2.7 are $\psi(\alpha) = \alpha^{1-\varepsilon}$, for any $\epsilon \in (0, 1]$.

It is instructive to compare these results with the picture provided by classical correlation. A straightforward computation shows that, if $(X_n)_n$ is a stationary ARCH(1) with $\mathbb{E}(\epsilon^2) = 1$ such that $a_1^2 \mathbb{E}(\epsilon^4) < 1$ (which implies existence of fourth moments), then the linear correlation of X_n^2 and X_{n+p}^2 exists and decays as a_1^p if $a_1 \to 0$. To quantify auto-dependence and heteroscedasticity in an ARCH if fourth moments do not exists one has to look for alternatives to linear correlation. The lower tail dependence coefficient $\lambda_{X_{n+p}|X_n}$ will always exist, and also converges to 0 as $a_1 \to 0$, by theorem 2.2 and the dominated convergence theorem - the exact order of convergence is less obvious, though. The generalized tail dependence coefficients $\lambda^{\psi}_{X_{n+k}|X_n}$ for ψ's such that $\psi(\alpha), \alpha/\psi(\alpha) \to 0$ will all be equal to 1/2, independent of a_1, as long as $a_1 > 0$. Similar remarks apply to non-stationary ARCH(1)'s.

It would be interesting to determine the precise behavior of $\lambda_{X_{n+p}|X_n}$ for a stationary ARCH(1) as either $a_1 \to 0$ or $p \to \infty$.

Theorems 2.6 and 2.7 are asymptotic results for α tending to 0, and it is reasonable to ask how indicative they are for what happens at a small but non-zero α. Taking for example $\alpha = 0.01$ and $\psi(\alpha) = \sqrt{\alpha}$, how close to 50% is the probability that X_{n+1} will be below its 10% quantile, given that X_n will have been below its 1%-one? In risk management terms, what is the probability that a portfolio will, on day $n+1$, incur a loss exceeding the 10% value-at-risk if on day n it will have suffered a loss greater than its 1% VaR? In section 7 we briefly investigate this question using Monte Carlo simulation. As we will see, depending on the value of a_1, this generalized lower tail dependence at say $\alpha = 0.01$ can be very close to its asymptotic value of 0.5, and even for values of a_1 as small as 0.1 the difference with the i.i.d. case (corresponding to an a_1 equal to 0) is significant.

Another interesting question for future research is whether empirical time series of financial returns behave like in theorems 2.2, 2.6 and 2.7. We note in this respect that these theorems provide an attractive quantitative reformulation of Mandelbrot's observation quoted in the first sentence of the introduction.

3. Lower tail dependence for stationary ARCH(1)

In this section we will prove theorems 2.2, 2.7 and 2.6. Let us put

$$F_{X_{n+p}}(x|X_n = y) = \mathbb{P}(X_{n+p} \leq x|X_n = y).$$

Lemma 3.1. *Given $y \in \mathbb{R}$ and $n, p \in \mathbb{N}$, $p \geq 1$, define functions $s_p(y;\cdot) : \mathbb{R}^{p-1} \to \mathbb{R}_{\geq 0}$, by*

$$s_p(y;\mathbf{x}) :=$$
$$(a_0 + a_0 a_1 x_{p-1}^2 + a_0 a_1^2 x_{p-2}^2 x_{p-1}^2 + \cdots + a_0 a_1^{p-1} x_1^2 \cdots x_{p-1}^2 + a_1^p x_1^2 \cdots x_{p-1}^2 y^2)^{1/2},$$

were $\mathbf{x} = (x_1, \cdots, x_{p-1})$. Then for all $n \geq 0$ and $p \geq 1$,

$$F_{X_{n+p}}(x|X_n = y) = \int_{\mathbb{R}^{p-1}} F_\epsilon\left(\frac{x}{s_p(y;\mathbf{x})}\right) \prod_{j=1}^{p-1} f_\epsilon(x_j) d\mathbf{x}. \tag{11}$$

Proof. Straightforward induction on p, observing that

$$\mathbb{P}(X_{n+1} \leq x|X_n = y) = F_\epsilon\left(\frac{x}{\sqrt{a_0 + a_1 y^2}}\right),$$

and using

$$\mathbb{P}(X_{n+p+1} \leq x|X_n = y) = \int_{\mathbb{R}} F_\epsilon\left(\frac{x}{\sqrt{a_0 + a_1 x_p^2}}\right) f_{X_{n+p}}(x_p|X_n = y)\, dx_p,$$

for the induction step. Here, $f_{X_{n+p}}(x|X_n = y)$ stands for the conditional density, $\frac{d}{dx} F'_{X_{n+p}}(x|X_n = y)$ (whose existence also follows by induction). QED

Corollary 3.2. *Let $x < 0$ be arbitrary. Then the function $y \to F_{X_{n+p}}(x|X_n = y)$ is decreasing on $\{y < 0\}$.*

Proof. Fix $x < 0$, and let $y \leq y' < 0$. Then $y^2 \geq y'^2$ and therefore $0 \leq s_p(y',\mathbf{x}) \leq s_p(y;\mathbf{x})$, for all $\mathbf{x} \in \mathbb{R}$. Since $x < 0$, it follows that $s_p(y';\mathbf{x})^{-1}x \leq s_p(y;\mathbf{x})^{-1}x$ and hence

$$F_\epsilon\left(\frac{x}{s_p(y';\mathbf{x})}\right) \leq F_\epsilon\left(\frac{x}{s_p(y;\mathbf{x})}\right).$$

Integration against $f_\epsilon(x_1) \cdots f_\epsilon(x_{p-1})$ then proves the corollary. QED

One can also give a proof by straightforward differentiation. Let $q_n(\alpha)$ be the α-th quantile of F_{X_n}.

Lemma 3.3. *Suppose that $q_n(\alpha) < 0$ and let $x < 0$. Then*

$$F_{X_{n+p}}(x|X_n = q_n(\alpha)) \leq \mathbb{P}(X_{n+p} \leq x|X_n \leq q_n(\alpha)) \leq \frac{1}{2}. \tag{12}$$

Proof. We have that

$$\mathbb{P}(X_{n+p} \leq x|X_n \leq q_n(\alpha)) = \frac{1}{\alpha} \int_{-\infty}^{q_n(\alpha)} F_{X_{n+p}}(x|X_n = y)\, f_{X_n}(y)\, dy, \tag{13}$$

where $f_{X_n} = F'_{X_n}$, the probability density. Integrating by parts, and using that $F_{X_n}(q_n(\alpha)) = \alpha$, we see that (13) equals

$$\begin{aligned} & F_{X_{n+p}}(x|X_n = q_n(\alpha)) - \int_{-\infty}^{q_n(\alpha)} \frac{d}{dy} F_{X_{n+p}}((x|X_n = y)\, \frac{F_{X_n}(y)}{\alpha}\, dy \\ \geq \ & F_{X_{n+p}}(x|X_n = q_n(\alpha)), \end{aligned}$$

by corollary 3.2. The other inequality in (3.3) follows from

$$\begin{aligned} \mathbb{P}(X_{n+p} \leq 0|X_n \leq q_n(\alpha)) &= \frac{1}{\alpha} \int_{-\infty}^{q_n(\alpha)} F_{X_{n+p}}(0|X_n = y)\, f_{X_n}(y)\, dy \\ &= \frac{1}{2}, \end{aligned}$$

by (11), the symmetry of ϵ and the definition of $q_n(\alpha)$. QED

Recall the definition of the twisted tail-dependence coefficients from the introduction. We then can state:

Corollary 3.4. *Let $\psi =: (0,1] \to (0,1]$ be defined in a neighborhood of 0 such that*

$$\lim_{\alpha \to 0} \frac{q_{n+p}(\psi(\alpha))}{q_n(\alpha)} = 0. \tag{14}$$

Then $\lambda^{\psi}_{X_{n+p}|X_n} = 1/2$.

Proof. We first observe that (14) implies that

$$\frac{q_{n+p}(\psi(\alpha))}{s_p(\mathbf{x}, q_n(\alpha))} \to 0, \quad \alpha \to 0.$$

By (11) and dominated convergence, $F_{X_{n+p}}(q_{n+p}(\alpha)|X_n = q_n(\alpha)) \to \frac{1}{2}$. The corollary now follows from (12). QED

Proof of theorem 2.7. If $(X_n)_{n \geq 0}$ is stationary, with $X_0 = X_\infty$, then $q_n(\alpha) = F^{\leftarrow}_{X_\infty}(\alpha) \simeq (c_\infty/\alpha)^{1/\kappa_\infty}$, as $\alpha \to 0$, for all n. It follows that (14) is equivalent to $\alpha = o(\psi(\alpha))$, $\alpha \to 0$. QED

Proof of theorem 2.2. Taking $x == q_n(\alpha) = F^{\leftarrow}_{X_\infty}(\alpha) =: q_\infty(\alpha)$ in lemma 3.1 (where X_∞ denotes the stationary ARCH(1)) and using that $q_{n+p}(\alpha) = q_\infty(\alpha)$ also, lemma 3.1 implies that

$$\lambda_{X_{n+p}|X_n}(\alpha) = \frac{1}{\alpha} \int_{-\infty}^{q_\infty(\alpha)} \int_{\mathbb{R}^{p-1}} F_\epsilon\left(\frac{q_\infty(\alpha)}{s_p(y; \mathbf{x})}\right) \prod_{j=1}^{p-1} f_\epsilon(x_j)\, d\mathbf{x}\, dF_\infty(y), \tag{15}$$

(compare (13), where $F_\infty(y) = F_{X_\infty}(y)$, the distribution function of X_∞. Integrating by parts with respect to y, and using that $F_\infty(q_\infty(\alpha)) = \alpha$, we find that

$$\lambda_{X_{n+p}|X_n}(\alpha) = \int_{\mathbb{R}^{p-1}} F_\epsilon\left(\frac{q_\infty(\alpha)}{s_p(q_\infty(\alpha);\mathbf{x})}\right) d\mathbf{x} \tag{16}$$

$$+\frac{q_\infty(\alpha)}{\alpha}\int_{-\infty}^{q_\infty(\alpha)}\int_{\mathbb{R}^{p-1}} f_\epsilon\left(\frac{q_\infty(\alpha)}{s_p(y;\mathbf{x})}\right)\frac{a_1^p x_1^2\cdots x_{p-1}^2 y}{s_p(y;\mathbf{x})^3}\prod_{j=1}^{p-1} f_\epsilon(x_j)F_\infty(y)\,d\mathbf{x}dy.$$

As $\alpha \to 0$, the integrand of the first term on the right tends to

$$F_\epsilon\left(\frac{q_\infty(\alpha)}{s_p(q_\infty(\alpha);\mathbf{x})}\right) \to F_\epsilon\left(\frac{-1}{a_1^{p/2}|x_1|\cdots|x_{p-1}|}\right).$$

We next change variables in the second integral on the right hand side of (16): $y = |q_\infty(\alpha)|z$. Since $F_\infty(x) \simeq c_\infty|x|^{-\kappa_\infty}$ and $|q_\infty(\alpha)| \simeq (c_\infty/\alpha)^{1/\kappa_\infty}$, it follows that

$$\alpha^{-1}F_\infty\left(|q_\infty(\alpha)|z\right) \to |z|^{-\kappa_\infty}, \quad \alpha \to 0.$$

Since $|q_\infty(\alpha)|^3/s_p(q_\infty(\alpha)z,\mathbf{x})^3 \to a_1^{-3p/2}|x_1|^{-3}\cdots|x_{p-1}|^{-3}|z|^{-3}$. We therefore find that, momentarily writing $A = a_1^{p/2}|x_1|\cdots|x_{p-1}|$,

$$\begin{aligned}
\lambda_{X_{n+p}|X_n} &= \int_{\mathbb{R}^{p-1}} F_\epsilon\left(-\frac{1}{A}\right)\prod_{j=1}^{p-1} f_\epsilon(x_j)\,d\mathbf{x} \\
&\quad -\int_{-\infty}^{-1}\int_{\mathbb{R}^{p-1}} f_\epsilon\left(-\frac{1}{A|z|}\right)\frac{1}{A}\frac{z}{|z|^3}\prod_{j=1}^{p-1} f_\epsilon(x_j)\,|z|^{-\kappa_\infty}d\mathbf{x}dz \\
&= \kappa_\infty\int_{-\infty}^{-1}\int_{\mathbb{R}^{p-1}} F_\epsilon\left(-\frac{1}{A|z|}\right)\prod_{j=1}^{p-1} f_\epsilon(x_j)\,|z|^{-\kappa_\infty-1}\,d\mathbf{x}dz,
\end{aligned}$$

after a second integration by parts, which proves (6). QED

Remark 3.5. The proof would simplify if we would know that X_∞ has a density $F_\infty'(x)$ with the differentiated asymptotics $\kappa_\infty c_\infty|x|^{-\kappa_\infty-1}$, $|x| \to \infty$, since then the partial integrations would no longer be necessary.

4. Asymptotic probabilities for non-stationary ARCH(1)'s

Let $(X_n)_{n\in\mathbb{N}}$ be the ARCH(1)-process (1), starting with an a.s. initial value $X_0 = x_0 \in \mathbb{R}$. We assume that f_ϵ is symmetric and satisfies condition 2.3: letting $\kappa' = \kappa_\epsilon + 1$, then

$$f_\epsilon(x) = \frac{c_\epsilon}{x^{\kappa'}} + \rho_\epsilon(x), \tag{17}$$

where

$$\left|\left(x\frac{d}{dx}\right)^\nu \rho_\epsilon(x)\right| \le \frac{C}{x^{\kappa'+\eta_0}}, \quad \nu = 0,1,2, \tag{18}$$

for suitable constants $\eta_0, C > 0$, for all $x > 0$. The successive realizations X_n of the ARCH(1) will then have a density f_{X_n} which, because of the Markov property of the ARCH(1), will be given by

$$f_{X_{n+1}}(x) = \int_{\mathbb{R}} f_\varepsilon\left(\frac{x}{\sqrt{a_0+a_1y^2}}\right) f_{X_n}(y)\frac{dy}{\sqrt{a_0+a_1y^2}}, \tag{19}$$

with $f_{X_1} = \sigma^{-1}f_\epsilon(x/\sigma_1)$, $\sigma_1 = \sqrt{a_0+a_1x_0^2}$. In this section we will prove the following truncated asymptotic expansion for the f_{X_n}; observe that, by symmetry, we can limit ourselves to $x > 0$.

Theorem 4.1. *With the above notations we have that*

$$f_{X_n}(x) = \frac{1}{\sigma_1 a_1^{(n-1)/2}} \varphi_n\left(\frac{x}{\sigma_1 a_1^{(n-1)/2}}\right), \tag{20}$$

with the function φ_n having a truncated expansion

$$\varphi_n(x) = \sum_{\nu=0}^{n-1} c_{\nu;n} \frac{(\log x)^\nu}{x^{\kappa'}} + r_n(x), \quad x > 0, \tag{21}$$

whose error $r_n(x)$ satisfies, for any positive $\eta < \eta_0$, an estimate

$$|r_n(x)| \leq \frac{C_\eta}{x^{\kappa'+\eta}}. \tag{22}$$

The coefficients $c_{\nu;n}$ and the remainder-function r_n may depend on a_0, a_1 and σ_1, but both the $c_{\nu;n}$ and the constants C_η in (22) will remain bounded as long as $\max(\sigma_1^{-1} a_1^{-1/2} a_0, \cdots, \sigma_1^{-1} a_1^{-(n-1)/2} a_0)$ *remains bounded. Moreover, the top order coefficient is given by*

$$c_{n-1;n} = 2^{n-1} \frac{c_\epsilon^n}{(n-1)!}, \tag{23}$$

and is therefore independent of a_0, a_1 and σ_1.

We obtain a corresponding result for the cumulative distribution function F_{X_n} by integration. We limit ourselves to the top-order asymptotics. Recalling that $\kappa = \kappa' - 1$, a simple integration by parts gives:

Corollary 4.2. *We have that*

$$F_{X_n}(x) = \Phi_n\left(\frac{x}{\sigma_1 a_1^{(n-1)/2}}\right), \tag{24}$$

where, as $x \to -\infty$,

$$\Phi_n(x) = C_n \frac{(\log|x|)^{n-1}}{|x|^\kappa} + R_n(x), \tag{25}$$

with $C_n = c_{n-1;n}/\kappa = 2^{n-1} c_\epsilon^n / \kappa(n-1)!$, and error $R_n(x)$ satisfying

$$|R_n(x)| \leq C \frac{|\log|x||^{n-2}}{|x|^\kappa}, \quad x < 0, \tag{26}$$

with uniformly bounded constant C whenever $\max(\sigma^{-1} a_1^{-1/2} a_0, \cdots, \sigma_1^{-1} a_1^{-(n-1)/2} a_0)$ *stays bounded.*

Remark 4.3. The asymptotics for x fixed, $n \to \infty$ are different from those for n fixed $|x| \to \infty$. It would be interesting to derive joint asymptotics in (x, n).

Theorem 4.1 will be proved by induction on n, based on (19). We will first analyze the case of $a_1 = \sigma_1 = 1$, to which the case of general a_0, a_1 and σ_1 will then be reduced.

4.1. **The case** $a_1 = 1$**.** We first take both a_1 and σ_1 equal to 1, and look for a truncated asymptotic expansion of f_{X_n} which is uniform in the parameter a_0 for bounded a_0: $a_0 \in [0,1]$ say (we might have let a_0 be restricted to any bounded interval in $[0,\infty)$). Define the operator F on $L^1([0,\infty))$ by

$$F(v)(x) = \int_0^\infty f_\varepsilon\left(\frac{x}{\sqrt{a_0+y^2}}\right) v(y) \frac{dy}{\sqrt{a_0+y^2}}; \tag{27}$$

F is a positivity-preserving operator on $L^1([0,\infty))$ of norm 1. Let $||\cdot||_1$ be the L^1-norm with respect to Lebesgue-measure on $[0,\infty)$.

Lemma 4.4. *Let $\{v(\cdot\ ; a_0)\ ; a_0 \geq 0\}$ be a family of non-negative functions in $C^1([0,\infty))$, such that $||v(\cdot\ ; a_0)||_1 = 1$, for all a_0, and such that $v(x; a_0)$ has the truncated asymptotic expansion*

$$v(x; a_0) = \sum_{\nu=0}^{n-1} c_\nu \frac{(\log x)^\nu}{x^{\kappa'}} + r(x), \quad x > 0, \tag{28}$$

with constants $c_\nu = c_\nu(v; a_0)$ uniformly bounded in $a_0 \leq 1$, and with remainder $r(x) = r(x; v, a_0)$ satisfying

$$\left|\left(x\frac{d}{dx}\right)^\nu r(x)\right| \leq \frac{C}{x^{\kappa'+\eta}}, \quad x > 0,\ \nu = 0, 1, \tag{29}$$

for suitable positive constants C and η with $\eta < 1$, uniformly for $a_0 \leq 1$. Let $u = u(\cdot; a_0)$ be given by

$$u(x; a_0) = F(v(\cdot\ ; a_0))(x).$$

Then $u \in C^1([0,\infty))$, $u \geq 0$, and u has a truncated expansion

$$u(x; a_0) = \sum_{\nu=0}^{n} c'_\nu \frac{(\log x)^\nu}{x^{\kappa'}} + r'(x), \quad x > 0, \tag{30}$$

with the coefficients $c'_\nu = c_\nu(u; a_0)$ again uniformly bounded in $a_0 \leq 1$ and $r'(x) = r'(x; v, a_0)$ satisfying estimates (29) with η replaced by any smaller $\eta' < \eta$, also uniformly in $a_0 \leq 1$. Moreover

$$c'_n = \frac{c_\epsilon c_{n-1}}{n}. \tag{31}$$

The truncated expansion (28) is compatible with $v(\cdot; a_0)$ being in L^1 since $\kappa' > 1$.

Proof. The idea is to write u as a Mellin convolution (see Appendix), by making the change of variables $z = \sqrt{y^2 + a_0}$. This gives

$$\begin{aligned} u(x) &= f_\epsilon * \widetilde{v}(x) \\ &= \int_0^\infty f_\epsilon\left(\frac{x}{z}\right) \widetilde{v}(z) \frac{dz}{z}, \end{aligned}$$

with

$$\widetilde{v}(z) = \frac{z}{\sqrt{z^2 - a_0}} v\left(\sqrt{z^2 - a_0}\right) \mathbf{1}_{[\sqrt{a_0},\infty)}(z), \tag{32}$$

$\mathbf{1}_A$ being the indicator function of a set A. However, in doing so we introduce a singularity at $z = \sqrt{a_0}$, which would spoil the decay properties of the Mellin transform needed later, and which we will therefore eliminate by a preliminary smooth cut-off. Let $\chi \in C^\infty([0,\infty))$ be such that $0 \leq \chi \leq 1$, supp $(\chi) \subset [A,\infty)$, $\chi = 1$ on $[2A,\infty)$, where $A > 0$ will be chosen below, and write $v = v(x; a_0)$ as

$$v = v_0 + v_1, \quad v_1 = \chi v,\ v_0 = (1-\chi)v. \tag{33}$$

We then study $u_0 := F(v_0)$ and $u_1 := F(v_1)$ separately.

First of all, by the hypothesis (17) on f_ϵ, we have that

$$F(v_0)(x) = \int_0^\infty f_\epsilon\left(\frac{x}{\sqrt{y^2+a_0}}\right) v_0(y)\frac{dy}{\sqrt{a_0+y^2}}$$
$$= x^{-\kappa'} c_\epsilon \int_0^\infty (y^2+a_0)^{\frac{\kappa'-1}{2}} v_0(y)dy + \int_0^\infty (y^2+a_0)^{-1/2}\rho_\epsilon\left((y^2+a_0)^{-1/2}x\right) v_0(y)dy$$
$$=: \frac{c_0'}{x^{\kappa'}} + r_0'(x),$$

with

$$c_0' = c_0'(a_0) := c_\epsilon \int_0^\infty (y^2+a_0)^{\frac{\kappa'-1}{2}} v_0(y)dy, \tag{34}$$

and

$$r_0'(x) = r_0'(x;a_0) = \int_0^\infty \rho_\epsilon\left(\frac{x}{\sqrt{y^2+a_0}}\right)\frac{dy}{\sqrt{y^2+a_0}}. \tag{35}$$

Since $\kappa' > 1$ and v_0 is supported in $[0, 2A]$, it follows that $c_0'(a_0)$ is uniformly bounded, for $a_0 \leq 1$, by

$$c_\epsilon(4A^2+1)^{\frac{\kappa'-1}{2}}||v_0||_1 = c_\epsilon(4A^2+1)^{\frac{\kappa'-1}{2}},$$

since $0 \leq (1-\chi) \leq 1$ and $||v||_1 = 1$.

Likewise, for $r_0'(x;a_0)$ we estimate, using (18) (with $\nu = 0$):

$$\begin{aligned} |r_0'(x;a_0)| &\leq \frac{C}{x^{\kappa'+\eta_0}}\int_0^A (y^2+a_0)^{\frac{\kappa'+\eta_0-1}{2}} v_0(y)dy \\ &\leq \frac{C'}{x^{\kappa'+\eta_0}}, \end{aligned}$$

with $C' = C(4A^2+a_0)^{(\kappa'+\eta_0-1)/2}$. This proves that $F(v_0)$ has a truncated asymptotic series of the desired form (30), (29).

We next turn to $u_1 = F(v_1)$. First rewrite u_1 as a Mellin transform, $u_1 = f_\epsilon * \widetilde{v}_1$, with $\widetilde{v}_1$ defined by (32) , with v replaced by v_1. Observe that if we take $A > 1 \geq a_0$, then $\widetilde{v}_1$ will be C^1 on $[0,\infty)$. From (28), (29), and remembering that $\eta < 1$, we easily obtain that

$$\begin{aligned} \widetilde{v}_1(x) &= (\chi\, v)\left(\sqrt{x^2-a_0}\right)\frac{x}{\sqrt{x^2-a_0}} \\ &= \sum_{\nu=0}^{n-1} c_\nu \frac{(\log x)^\nu}{x^{\kappa'}} + \widetilde{r}_1(x), \end{aligned}$$

with $\widetilde{r}_1 = \widetilde{r}_1(x;a_0)$ satisfying (29), uniformly in $a_0 \leq 1$ (with the same η but possibly a different C). It follows that the Mellin transform $\widetilde{v}_1^\# = \widetilde{v}_1^\#$ is holomorphic on $\{s \in \mathbb{C} : -\kappa' < \text{Re } s < 0\}$, and meromorphic on $\{s : -\kappa' - \eta < \text{Re } (s) < 0\}$ with

$$\widetilde{v}_1^\# = \sum_{\nu=1}^{n} \frac{a_{-\nu}}{(s+\kappa')^\nu} + h(s), \tag{36}$$

with $a_{-\nu} = (\nu-1)!c_{\nu-1}$ and $h = h(s)$ holomorphic on $\{s : -\kappa' - \eta < \text{Re } (s) < 0\}$ (cf. the Appendix). It is easily verified that (28) and (29), together with the differentiability of v, imply that the truncated series (28) can be differentiated, and that

$$x\frac{d}{dx}\widetilde{v}_1(x) = \sum_{\nu=0}^{n-1} c_\nu'' \frac{(\log x)^\nu}{x^{\kappa'}} + x\frac{d}{dx}\widetilde{r}_1(x),$$

for suitable constants c_ν'', which are simply obtained by term-by term differentiation of the truncated expansion for $\widetilde{v}_1$. (A similar remark applies to (17) and (18).) Taking Mellin transforms, we conclude that, given that $\widetilde{r}_1$ satisfies (29) with $\nu = 1$,

$$s\widetilde{v}_1^{\#}(s) = \sum_{\nu=0}^{n-1} \frac{A_\nu}{(s+\kappa')^\nu} + g(s),$$

with $g = g(s)$ holomorphic on $\{s : -\kappa' - \eta < \text{Re }(s) < 0\}$, and bounded on each closed sub-strip $\{s : -\kappa' - \eta + \varepsilon \leq \text{Re }(s) \leq -\varepsilon\}$, $\varepsilon > 0$. From this we conclude that

(i) If $\sigma \in (-\kappa' - \eta, 0)$ and $\sigma \neq -\kappa'$, then

$$t \to \widetilde{v}_1^{\#}(\sigma + it) \ \in \ L^2(\mathbb{R}), \tag{37}$$

and

(ii) For any closed sub-interval $[a, b] \subset (-\kappa' - \eta, 0)$, we have that

$$\max_{\sigma \in [a,b]} \left| \widetilde{v}_1^{\#}(\sigma \pm iR) \right| \to 0 \ \text{ as } R \to \infty. \tag{38}$$

Similar assertions can be proved, in the same way, for $f^{\#}(s)$, starting from (17) and (18) with $\nu = 0$ and 1. It follows that for arbitrary $\rho \in (-\kappa', 0)$ we can apply the inversion formula for the Mellin transform to write $u_1(x) = F(v_1)(x)$ as

$$u_1(x) = \frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} f^{\#}(s)\widetilde{v}_1^{\#}(s) x^s ds, \tag{39}$$

(the integral converges since the product $f^{\#}(s)\widetilde{v}_1^{\#}(s)$ is in L^1 of the line $\{\text{Re }(s) = \sigma\}$). Moreover, the right conditions are met to shift the integration path to $\{\text{Re } s = -\kappa - \eta'\}$, where $\eta' < \eta$ is arbitrary. In doing so, we will pick up the residue in $s = -\kappa'$:

$$\begin{aligned} &\frac{1}{2\pi i}\text{Res }_{s=-\kappa'} \left(\left(\frac{c_\epsilon}{s+\kappa'} \right) \left(\sum_{\nu=1}^{n} \frac{(\nu-1)! c_{\nu-1}}{(s+\kappa')^\nu} + h(s) \right) x^s \right) \\ &= \sum_{\nu=0}^{n} c_\nu' \frac{(\log x)^\nu}{x^{\kappa'}}, \end{aligned} \tag{40}$$

as one sees by writing $x^s = x^{-\kappa'} \exp\left((s+\kappa')\log x\right)$. It follows in particular that

$$c_n' = \frac{c_\epsilon c_{n-1}}{n}. \tag{41}$$

Hence

$$u_1(x) = \sum_{\nu=0}^{n} c_\nu' \frac{(\log x)^\nu}{x^{\kappa'}} + r'(x),$$

where the remainder term is given by the line-integral over Re $s = -\kappa - \eta'$:

$$r'(x) = \frac{1}{2\pi i} \int_{\text{Re } s = -\kappa - \eta'} f^{\#}(s)\widetilde{v}_1^{\#}(s) x^s ds,$$

which, by an obvious estimate, will satisfy

$$|r'(x)| \leq C / x^{\kappa' + \eta'}. \tag{42}$$

To show that $x(d/dx)r'$ satisfies the same estimate, we first differentiate u_1, to find

$$x\frac{d}{dx}u_1 = \left(x\frac{df_\epsilon}{dx} \right) * \widetilde{v}_1.$$

Now since xdf_ε/dx satisfies (17) and (18) for $\nu = 0, 1$, we can repeat the argument above to conclude that there exist constants c''_ν such that

$$x\frac{d}{dx}u_1 = \sum_{\nu=0}^{c} c''_\nu \frac{(\log x)^\nu}{x^{\kappa'}} + r''(x),$$

with $r''(x)$ satisfying an estimate of the type (42). That is, $x(d/dx)u_1$ will have a truncated asymptotic series of the same type as u_1. A moment's thought then shows that, since the differential operator xd/dx, formally applied to the expansion of u_1, yields an expansion of the same form, we necessarily must have that $r'' = (xd/dx)r'$. This shows that r' satisfies (29) (with η' instead of η), and completes the proof of the Main Lemma. QED

4.2. **Proof of theorem 4.1.** Lemma 4.4 allows us to quickly prove theorem 4.1, using induction on n. Fix an initial value $x_0 \in \mathbb{R}$ and put $\sigma_1 = \sqrt{a_0 + a_1 x_0^2}$. Our induction hypothesis is that $f_{X_n}(x)$ is of the form (20), with $\varphi_n(x)$ having a truncated expansion (21), with the $c_{\nu;n}$ and r_n satisfying the stated uniformity conditions. For $n = 1$, this is satisfied with $\varphi_1 = f_\epsilon$, by the hypothesis on f_ϵ.

Let F_{a_0,a_1} be the norm-preserving positive operator on $L^1(\mathbb{R}_{\geq 0}$ defined by

$$F_{a_0,a_1}(v)(x) = 2\int_0^\infty f_\varepsilon\left(\frac{x}{\sqrt{a_0 + a_1 y^2}}\right) v(y) \frac{dy}{\sqrt{a_0 + a_1 y^2}}, \tag{43}$$

so that the F defined by (27) equals $F_{a_0,1}$. Then we have, by the Markov property of our ARCH(1)-process, and the symmetry of f_ϵ, that

$$\begin{aligned} f_{X_{n+1}}(x) &= 2\, F_{a_0,a_1}\left(f_{X_n}\right)(x) \\ &= 2\, F_{a_0,a_1}\left(\frac{1}{\sigma_1 a_1^{(n-1)/2}}\varphi_n\left(\frac{y}{\sigma_1 a_1^{(n-1)/2}}\right)\right)(x) \\ &= \frac{2}{\sigma_1 a_1^{n/2}}\, F_{\sigma^{-1}a_1^{-n/2}a_0,1}\left(\varphi_n\right)\left(\frac{x}{\sigma_1 a_1^{n/2}}\right), \end{aligned}$$

by an easy change of variables in the integral defining F_{a_0,a_1}. Put

$$\varphi_{n+1} := F_{\sigma^{-1}a_1^{-n/2}a_0,1}(\varphi_n).$$

Then by lemma 4.4, applied to $v = \varphi_n$, we find that

$$\varphi_{n+1}(x) = \sum_{\nu=0}^{n} c_{\nu;n+1}\frac{(\log x)^\nu}{x^{\kappa'}} + r_{n+1}(x),$$

with r_{n+1} satisfying (29), uniformly in $\max(\sigma_1^{-1}a_1^{-1/2}a_0, \cdots, \sigma_1^{-1}a_1^{-n/2}a_0)$, and uniformly bounded coefficients $c_{\nu;n+1}$ for these values of the parameters. Moreover, $c_{n;n+1} = 2c_\epsilon c_{n-1;n}/n$ and therefore, since $c_{0;1} = c_\varepsilon$, $c_{n-1;n} = 2^{n-1}c_\varepsilon^n/(n-1)!$. QED

5. Asymptotic behavior of the non-stationary VaR

For the computation of the $\alpha \to 0$ asymptotics of the tail-dependence function $\lambda^{x_0}_{X_{n+p}|X_n}(\alpha)$ we need to know the asymptotic behavior of the quantile functions $q^{x_0}_{X_n}(\alpha) = F^{x_0,\leftarrow}_{X_n}(\alpha)$. This can be computed from corollary 4.2:

Lemma 5.1. *As $\alpha \to 0$,*

$$q^{x_0}_{X_n}(\alpha) \simeq \sigma_1 a_1^{(n-1)/2}\left(\frac{C_n}{\alpha}\right)^{1/\kappa}\left(\frac{\log(\alpha^{-1}C_n)}{\kappa}\right)^{\frac{n-1}{\kappa}}, \tag{44}$$

in the usual sense of the quotient of the two sides tending to 1.

Proof. Since it is not true in general that the map $F \to F^{\leftarrow}$ is continuous with respect to the uniform topologies, we proceed cautiously.

First of all, since $F^{x_0}_{X_n}(x) = \Phi_n(x/\sigma_1 a_1^{(n-1)/2})$, we immediately have that

$$q^{x_0}_{X_n}(\alpha) = \sigma_1 a_1^{(n-1)/2}\Phi_n^{\leftarrow}(\alpha). \tag{45}$$

From corollary 4.2 we see that for any $\varepsilon > 0$ we can find a positive real number $R(\varepsilon)$ such that if $x < -R(\varepsilon)$, then

$$(C_n - \varepsilon)\frac{(\log|x|)^{n-1}}{|x|^{\kappa}} \le \Phi_{X_n}(x) \le (C_n + \varepsilon)\frac{(\log|x|)^{n-1}}{|x|^{\kappa}}.$$

Call the two functions on the left and right of this inequality $\Phi_n^{\pm\varepsilon}$. Since these are strictly increasing, and since $\Phi_n^{-\varepsilon}(x) \le \Phi_n(x) \le \Phi_n^{\varepsilon}(x)$, we will have that

$$(\Phi_n^{\varepsilon})^{-1}(\alpha) \le \Phi_n^{\leftarrow}(\alpha) \le (\Phi_n^{-\varepsilon})^{-1}(\alpha),$$

at least for those α for which the biggest of these three numbers is $\le -R(\varepsilon)$. (By theorem 4.1, $\Phi_n'(x) = \varphi_n(x)$ is strictly positive if x is sufficiently small, so that we could have replaced $\Phi_n^{\leftarrow}$ by its ordinary inverse also). To compute $(\Phi_n^{\varepsilon})^{-1}(\alpha)$ we use the following elementary lemma.

Lemma 5.2. *Let $a > 0$ and let $x = x(a)$ be the (unique) positive solution to*

$$\frac{(\log x)^{n-1}}{x^{\kappa}} = a.$$

Then

$$x(a) = a^{-1/\kappa}\left(\frac{\log a^{-1}}{\kappa}\right)^{\frac{n-1}{\kappa}} g(a),$$

where $g(a) \to 1$ as $a \to 0$.

Proof. Put $\log x = y$, $x = x(a)$. Then y has to solve

$$y - \frac{n-1}{\kappa}\log y = \log(a^{-1/\kappa}) =: A.$$

If we try a solution of the form $y = A + \frac{n-1}{\kappa}\log A + y'$, then $y' = y'(A)$ will have to solve:

$$\begin{aligned} y' &= \frac{n-1}{\kappa}\log\left(1 + \frac{n-1}{\kappa}\frac{\log A + y'}{A}\right) \\ &\le \left(\frac{n-1}{\kappa}\right)^2\left(\frac{\log A + y'}{A}\right). \end{aligned}$$

Hence we find that[1]

$$\begin{aligned} 0 &< y' \le \left(\frac{n-1}{\kappa}\right)^2 \frac{\log A/A}{\left(1 - \left(\frac{n-1}{\kappa}\right)^2 A^{-1}\right)} \\ &\le 2\left(\frac{n-1}{\kappa}\right)^2\frac{\log A}{A}, \end{aligned}$$

provided that $A \ge 2\kappa^{-2}(n-1)^2$. Exponentiating, and putting $g(a) := \exp y'(A)$, we find that

$$x(a) = a^{-1/\kappa}\left(\frac{\log a^{-1}}{\kappa}\right)^{(n-1)/\kappa} g(a),$$

where

$$1 \le g(a) \le \exp\left(2(n-1)^2\log A/\kappa^2 A\right), \tag{46}$$

[1]observe that if $y_0 := A + (n-1)\log A/\kappa$, then $y_0 - (n-1)\log y_0/\kappa \le A$, so that y' will have to be strictly positive

if $\log a^{-1/\kappa} \geq 2\kappa^{-2}(n-1)^2$. Hence $g(a) \to 1$ as $a \to 0$. QED

Returning to the proof of lemma 5.1, it follows that

$$(\Phi_n^{\pm\varepsilon})^{-1} = \left(\frac{(C_n \pm \varepsilon)}{\alpha}\right)^{1/\kappa} \left(\frac{\log(\alpha^{-1}(C_n \pm \varepsilon))}{\kappa}\right)^{(n-1)/\kappa} g_{\pm\varepsilon}(\alpha),$$

with $g_{\pm\varepsilon}(\alpha) \to 1$ as $\alpha \to 0$, uniformly in $0 \leq \varepsilon \leq \varepsilon_0 < C_n$ (as a glance at (46) shows). Putting

$$Q_n(\alpha) = \left(\frac{C_n}{\alpha}\right)^{1/\kappa} \left(\frac{\log(\alpha^{-1}C_n)}{\kappa}\right)^{(n-1)/\kappa},$$

we see that

$$\begin{aligned} &\left(\frac{C_n - \varepsilon}{C_n}\right)\left(\frac{\log(\alpha^{-1}C_n - \varepsilon)}{\log(\alpha^{-1}C_n)}\right)^{(n-1)/\kappa} g_{-\varepsilon}(\alpha) \\ \leq &\frac{\Phi_n^{\leftarrow}(\alpha)}{Q_n(\alpha)} \leq \left(\frac{C_n + \varepsilon}{C_n}\right)\left(\frac{\log(\alpha^{-1}C_n + \varepsilon)}{\log(\alpha^{-1}C_n)}\right)^{(n-1)/\kappa} g_{\varepsilon}(\alpha). \end{aligned}$$

Letting $\alpha \to 0$, we conclude that for all sufficiently small $\varepsilon > 0$,

$$(1 - \frac{\varepsilon}{C_n}) \leq \liminf_{\alpha\to 0} \frac{\Phi_n^{\leftarrow}(\alpha)}{Q_n(\alpha)} \leq \limsup_{\alpha\to 0} \frac{\Phi_n^{\leftarrow}(\alpha)}{Q_n(\alpha)} \leq (1 + \frac{\varepsilon}{C_n}),$$

and the lemma follows by letting $\varepsilon \to 0$. QED

6. Asymptotics of the non-stationary tail-dependence function

The asymptotic behavior of the conditional lower tail dependence function can be determined quite precisely, and turns out to be *universal* for the class of ARCH(1)'s whose innovations $(\epsilon_n)_n$ satisfy condition 2.3 with a given κ_ϵ and c_ϵ, in the sense that it does not depend on either a_0, a_1 or the initial value x_0, nor on the further details of f_ϵ:

Theorem 6.1. *Let $(X_n)_{n\in\mathbb{N}}$ be an ARCH(1) with a.s. initial value $X_0 = x_0$ and $a_1 > 0$ strictly positive. Suppose that ϵ_n has a twice differentiable symmetric probability density satisfying condition 2.3. Then for all $n, p \geq 1$,*

$$\lambda^{x_0}_{X_{n+p}|X_n}(\alpha) \simeq \gamma_{n,p} \frac{\left(\log(\log\alpha^{-1})\right)^p}{(\log\alpha^{-1})^p}, \quad \alpha \to 0, \tag{47}$$

where

$$\gamma_{n,p} = \frac{2^{2n+2p-3}}{p!(n-1)!(n+p-1)!} \frac{c_\epsilon^{2n+2p}}{\kappa_\epsilon^2}. \tag{48}$$

In particular, $\lambda^{x_0}_{X_{n+p}|X_n} = lim_{\alpha\to 0}\lambda^{x_0}_{X_{n+p}|X_n}(\alpha) = 0$.

Remarks 6.2. (i) Observe that the conditional lower tail dependence function (47) tends to 0 at an exponentially slower rate than for independent random variables. For values of a_1 such that the ARCH(1) has a stationary solution, the different conclusions of theorems 2.2 and 6.1 may appear strange, in view of the geometric ergodicity of the process (cf. [4] and its references). The explanation is that, in case of a stationary process, $q_{X_n}(\alpha) = q_{X_\infty}(\alpha) = q_{X_{n+p}}(\alpha)$ for all n and p, while for a conditional ARCH(1) and $n, p \geq 1$, $q_{X_{n+p}}(\alpha)/q_{X_n}(\alpha) \to \infty$ as $\alpha \to 0$ (although slowly), as follows from lemma 5.1.

(ii) The results of [9] can be used to derive the asymptotics of $\lambda^{x_0}_{X_{n+p}|X_n}$ in case of normally distributed ϵ_n. These have a qualitatively different behavior.

Proof of theorem 6.1 To simplify notations, we will write $q_n(\alpha)$ for $q_{X_n}^{x_0}(\alpha)$, and we will generally leave off the super-index x_0. Since the joint probability density of (X_{n+p}, X_n) is a product of the conditional density of $X_{n+p}|X_n$ and the density of X_n, we see as before that

$$\begin{aligned}
&\mathbb{P}\left(X_{n+p} \leq q_{n+p}(\alpha) | X_n \leq q_n(\alpha)\right) \\
&= \frac{1}{\alpha} \int_{-\infty}^{q_n(\alpha)} F_{X_{n+p}}(q_{n+p}(\alpha) | X_n = y)\, f_{X_n}(y)\, dy \\
&= \frac{1}{\alpha} \int_{-\infty}^{q_n(\alpha)} \Phi_p \left(\frac{q_{n+p}(\alpha)}{a_1^{(p-1)/2} \sqrt{a_0 + a_1 y^2}} \right) \varphi_n \left(\frac{y}{\sigma_1 a_1^{(n-1)/2}} \right) \frac{dy}{\sigma_1 a_1^{(n-1)/2}}, \\
&= \frac{1}{\alpha} \int_{1}^{\infty} \Phi_p \left(- \frac{q_{n+p}(\alpha)/q_n(\alpha)}{a_1^{(p-1)/2} \sqrt{a_0 + a_1 y^2}} \right) \varphi_n \left(\frac{q_n(\alpha) y}{\sigma_1 a_1^{(n-1)/2}} \right) \frac{q_n(\alpha)}{\sigma_1 a_1^{(n-1)/2}}\, dy,
\end{aligned}$$

where we have used (24) for $F_{X_{n+p}}(\cdot | X_n = y)$, with n instead of p and $X_n = y$ instead $X_0 = x_0$. Since $q_n(\alpha)/\sigma_1 a_1^{(n-1)/2} = \Phi_n^{\leftarrow}(\alpha)$, we have that

$$\frac{q_{n+p}(\alpha)/q_n(\alpha)}{a_1^{(p-1)/2} \sqrt{a_0 + a_1 y^2}} = \frac{\Phi_{n+p}^{\leftarrow}(\alpha)/\Phi_n^{\leftarrow}(\alpha)}{\sqrt{\widetilde{a}_0 + y^2}},$$

where we have put $\widetilde{a}_0 := a_1^{-1} a_0$. In what follows we will, without further comment, replace various quantities by their asymptotic equivalents: this can be made rigorous by standard estimates. First of all, by theorem 4.1,

$$\begin{aligned}
&\mathbb{P}\left(X_{n+p} \leq q_{n+p}(\alpha) | X_n \leq q_n(\alpha)\right) \qquad (49)\\
&\simeq \frac{c_{n-1;n}}{\alpha} \left(\Phi_n^{\leftarrow}(\alpha)\right)^{-\kappa} \int_1^\infty \Phi_p \left(\frac{\Phi_{n+p}^{\leftarrow}(\alpha)/\Phi_n^{\leftarrow}(\alpha)}{\sqrt{\widetilde{a}_0 + y^2}} \right) \frac{\left(\log \Phi_n^{\leftarrow}(\alpha) y\right)^{n-1}}{y^{\kappa+1}}\, dy.
\end{aligned}$$

By lemma 5.1,

$$\left(\Phi_n^{\leftarrow}\right)(\alpha))^{-\kappa} \simeq \frac{\alpha}{C_n} \left(\frac{\log \alpha^{-1} C_n}{\kappa} \right)^{-(n-1)},$$

where the α will cancel the $1/\alpha$ in front of (49). To simplify notations, we put $\lambda := \alpha^{-1} C_n$ and $\Lambda := \log \Phi_n^{\leftarrow}(\alpha) \simeq \kappa^{-1} \log \lambda$. Then, using lemma 5.1 once more, we see that

$$\Phi_{n+p}^{\leftarrow}(\alpha)/\Phi_n^{\leftarrow}(\alpha) \simeq \gamma (\log \lambda)^{p/\kappa}, \quad \gamma = \gamma_{n,p} := (C_n^{-1} C_{n+p})^{1/\kappa}$$

and (49) equals

$$\begin{aligned}
&\frac{c_{n-1;n}}{C_n} \left(\frac{\kappa}{\log \lambda} \right)^{n-1} \sum_{\nu=0}^{n-1} \binom{n-1}{\nu} \cdot \qquad (50)\\
&\Lambda^{n-1-\nu} \int_1^\infty \Phi_p \left(- \frac{\gamma (\log \lambda)^{p/\kappa}}{\sqrt{\widetilde{a}_0 + y^2}} \right) \frac{(\log y)^\nu}{y^\kappa} \frac{dy}{y}.
\end{aligned}$$

If we make the change of variables $z = \sqrt{\widetilde{a}_0 + y^2}$, we recognize the integrals as being the Mellin convolution, evaluated in $\gamma(\log \lambda)^{p/\kappa}$), of $\Phi_p(z)\, \mathbf{1}_{(-\infty,0)}$ with

$$g_\nu(z) := \frac{(\log \sqrt{z^2 - \widetilde{a}_0}\,)^\nu}{(z^2 - \widetilde{a}_0)^{(\kappa+1)/2}} \frac{z}{\sqrt{z^2 - \widetilde{a}_0}} \mathbf{1}_{[\widetilde{a}_0 + 1, \infty)}(z),$$

which asymptotically still equals $(\log z)^\nu / z^\kappa$. By the analysis of section 2, we know that the Mellin transform of $\Phi_p\ \mathbf{1}_{(-\infty,0)}$ is meromorphic on a strip $\{-\kappa - \eta <$

Re $s < 0\}$, $0 < \eta < 1$, with a unique pole of order p in $s = -\kappa$ and principal part $\sum_{\nu=0}^{p} a_{-\nu}(s+\kappa)^{-\nu}$, with $a_p = (p-1)!\, C_p$. The Mellin transform of g_ν equals

$$\begin{aligned} g_\nu^{\#}(s) &= \int_{\widetilde{a}_0+1}^{\infty} \frac{(\log y)^\nu}{y^\kappa}(y^2+\widetilde{a}_0)^{-s/2}\frac{dy}{y} \\ &= \int_1^\infty \frac{(\log y)^\nu}{y^\kappa}\, y^{-s}\left(1+\frac{\widetilde{a}_0}{y^2}\right)^{s/2}\frac{dy}{y} \\ &= \frac{\nu!}{(s+\kappa)^{\nu+1}} + H(s), \end{aligned}$$

where

$$H(s) = \int_1^\infty \left(\left(1+\frac{\widetilde{a}_0}{y^2}\right)^{s/2} - 1\right)\frac{(\log y)^\nu}{y^{\kappa+s+1}}dy,$$

is holomorphic on $\{-\kappa - 1 < \mathrm{Re} s < 0\}$, as follows easily form the asymptotic behavior of $(1+\widetilde{a}_0 y^{-2})^{s/2} - 1$ as $y \to \infty$. Using the arguments of section 3, one finds that

$$\begin{aligned} & \Lambda^{n-1-\nu}\cdot \int_1^\infty \Phi_p\left(-\frac{\gamma(\log\lambda)^{p/\kappa}}{\sqrt{\widetilde{a}_0+y^2}}\right)\frac{(\log y)^\nu}{y^\kappa}\,\frac{dy}{y} \\ \simeq\ & \Lambda^{n-1-\nu}\cdot \mathrm{Const}_\nu \cdot \frac{\left(\log(\gamma(\log\lambda)^{p/\kappa})\right)^{p+\nu}}{\gamma^\kappa(\log\lambda)^p} \\ =\ & O\left((\log\lambda)^{n-1-\nu}\cdot\frac{(\log(\log\lambda))^{p+\nu}}{(\log\lambda)^p}\right). \end{aligned}$$

It follows that the dominant term in (50) is the one with $\nu = 0$. Computing the constant Const_0, which gives C_p/p, we find that (49) is asymptotically equivalent to

$$\begin{aligned} & \frac{c_{n-1;n}\, C_p}{p\, C_n}\left(\frac{\kappa}{\log\lambda}\right)^{n-1}\Lambda^{n-1}\frac{(\log\log\lambda)^p}{(C_{n+p}/C_n)(\log\lambda)^p} \\ \simeq\ & \frac{c_{n-1;n}\, C_pC_{p+n}}{p}\,\frac{(\log\log\alpha^{-1})^{n-1}}{(\log\alpha^{-1})^p}. \end{aligned}$$

This proves theorem 6.1. QED

Proof of theorem 2.4. Immediate from (47). QED

Proof of theorem 2.6. We may use lemma 3.3 and corollary 3.4., with $\mathbb{P}$, $q_n(\alpha)$ and $\lambda_{X|Y}^{\psi}$ replaced by $\mathbb{P}^{x_0}$, $q_n^{x_0}(\alpha)$, and $\lambda_{X_{n+p}|X_n}^{\psi,x_0}$. By lemma 5.1, (44), we see that, modulo an immaterial constant,

$$\frac{q_{n+p}^{x_0}(\psi(\alpha))}{q_n^{x_0}(\alpha)} \simeq \left(\frac{\alpha}{\psi(\alpha)}\right)^{1/\kappa}\frac{\left(\log\psi(\alpha)^{-1}C_{n+p}\right)^{(n+p-1)/\kappa}}{\left(\log\alpha^{-1}C_n\right)^{(n-1)/\kappa}},$$

for $\alpha \to 0$. If $\psi(\alpha) \to 0$ and $\alpha\,(\log\psi(\alpha)^{-1})^p/\psi(\alpha) \to 0$, then certainly $\alpha/\psi(\alpha) \to 0$ and consequently $\log\psi(\alpha)^{-1} = O(\log\alpha^{-1})$ (using that $\psi(\alpha) < 1$ for sufficiently small α). It follows that the expression on the right hand side will be dominated by a constant times $\left(\alpha(\log\psi(\alpha)^{-1})^p/\psi(\alpha)\right)^{1/\kappa}$ and hence tends to 0 with α. QED

7. Generalized lower tail dependence functions at non-zero α

Consider a stationary ARCH(1) $(X_n)_{n\geq 0}$ and let

$$\lambda^\psi(\alpha) := \mathbb{P}\left(X_{n+p} < q_{X_{n+p}}(\psi(\alpha)) | X_n < q_{X_n}(\alpha)\right), \tag{51}$$

which may be called a generalized lower tail dependence function. Theorems 2.6 and 2.7 show that the limit of $\lambda^\psi(\alpha, p)$ as $\alpha \to 0$ is equal to $1/2$, under appropriate conditions on ψ. To see what happens at a small but non-zero α we have computed some $\lambda^{\psi}(\alpha, p)$ using Monte Carlo simulations. What follows is not intended as a complete simulation study into the behavior of these functions, but as a simple exploratory investigation. Its main aim is to illustrate that our theorems are relevant at values of α equal to 0.01 or 0.05, which are the values typically used in risk management. We have not, in particular, tried to improve the precision by using any variance reduction or importance sampling tecniques, leaving these for future work. We will also limit ourselves to $\psi(\alpha) = \sqrt{\alpha}$, as a typical example of a function ψ satisfying the conditions in theorems 2.6 and 2.7.

It is important to note that what is important at non-zero α is not the conditional tail probability (51) itself, but rather its difference with $\psi(\alpha)$, the tail probability for the independent case. Remembering that we took $\psi(\alpha) = \sqrt{\alpha}$, we therefore introduce the *(generalized) excess lower tail probability*

$$e(\alpha, p) := \lambda^{\sqrt{\cdot}}(\alpha, p) - \sqrt{\alpha}. \tag{52}$$

We first look at stationary processes. Taking $a_0 = 0.001$, we have simulated 250 ARCH(1)'s of length 10^4 with Student t_4 innovations ϵ_n and a_1 ranging from 0 to 1.6 with step-size 0.1. We note that such an ARCH(1) is strictly stationary if $a_1 < \sqrt{e} \simeq 1.6489$. For each Monte Carlo run we computed the empirical conditional probabilities (51) and the difference (52), and taken their averages over the different runs, to be denoted by $\hat{\lambda}^{\sqrt{\cdot}}(\alpha, p)$, $\hat{e}(\alpha, p)$, as well as their standard deviation of, to assess the significance of the estimated values. The results, rounded off to two significant figures, are given in the following table.

a_1	$\hat{\lambda}^{\sqrt{\cdot}}(0.05, 1)$	$\hat{e}(0.05, 1)$	std		$\hat{\lambda}^{\sqrt{\cdot}}(0.01, 1)$	$\hat{e}(0.01, 1)$	std
0	0.23	0	0.02		0.1	0	0.03
0.1	0.29	0.06	0.02		0.23	0.13	0.04
0.2	0.32	0.1	0.02		0.3	0.2	0.05
0.3	0.35	0.13	0.02		0.34	0.24	0.05
0.4	0.37	0.15	0.02		0.38	0.28	0.05
0.5	0.39	0.17	0.02		0.40	0.30	0.05
0.6	0.41	0.19	0.02		0.43	0.33	0.05
0.7	0.43	0.20	0.02		0.44	0.34	0.05
0.8	0.44	0.22	0.02		0.45	0.35	0.05
0.9	0.45	0.23	0.02		0.47	0.37	0.05
1.0	0.47	0.24	0.02		0.48	0.38	0.05
1.1	0.47	0.25	0.02		0.49	0.39	0.05
1.2	0.48	0.26	0.02		0.49	0.39	0.05
1.3	0.49	0.26	0.02		0.49	0.39	0.05
1.4	0.49	0.27	0.02		0.5	0.4	0.05
1.5	0.49	0.27	0.02		0.49	0.39	0.05
1.6	0.49	0.27	0.02		0.5	0.4	0.05

Table 7.1 - Stationary tail probabilities as function of a_1.

The table shows for example that for a value of $a_1 = 0.5$ there is a close to 40% change that X_{n+1} will exceed its 95% VaR, given that X_n will have done so. This conditional probability is about the same for the 99% VaR, and in that case about 30% bigger than what it would have been had succesive returns been independent. The effect becomes more pronounced with increasing values of a_1.

The next table gives estimated excess tail probabilities $\hat{e}(p) := \hat{e}(0.01, p)$ for lags p varying from 1 to 10, at an $\alpha = 1\%$ and for three values of a_1, namely $a_1 = 0.5, 1$ and 1.5. We take $a_0 = 0.001$ and $\epsilon \sim t_4$, as before.

p	$\hat{e}(p)$(std), $a_1 = 0.5$	$\hat{e}(p)$(std), $a_1 = 1$	$\hat{e}(p)$(std), $a_1 = 1.5$
1	0.31 (0.05)	0.38 (0.05)	0.40 (0.05)
2	0.20 (0.05)	0.35 (0.06)	0.39 (0.05)
3	0.13 (0.05)	0.30 (0.06)	0.38 (0.05)
4	0.07 (0.04)	0.26 (0.06)	0.37 (0.06)
5	0.05 (0.04)	0.22 (0.06)	0.36 (0.05)
6	0.03 (0.04)	0.19 (0.06)	0.35 (0.06)
7	0.02 (0.03)	0.16 (0.06)	0.34 (0.06)
8	0.01 (0.04)	0.14 (0.06)	0.32 (0.06)
9	0.01 (0.03)	0.11 (0.05)	0.31 (0.07)
10	0 (0.03)	0.10 (0.05)	0.29 (0.07)

Table 7.2 - Stationary excess tail probabilities as function of the lag p.

As was to be expected, the excess tail probabilities decrease with p. This decrease is slower the larger a_1 is (and the more pronounced therefore the ARCH-effect): for $a_1 = 0.5$, our estimate of $e(0.01, p)$ is no longer significantly different from 0 from $p = 6$ onwards, while for $a_1 = 1.5$, we found for example that $\hat{e}(0.01, p = 20) = 0.16(0.08)$, which indicates a long persistence of return shocks at time 1.

We have repeated these computations for non-stationary ARCH(1)'s with a.s. initial value $X_0 = x_0$. Let

$$\lambda^{\sqrt{\cdot};x_0}(\alpha; k, k+p) := \mathbb{P}^{x_0}\left(X_{n+p} < q_{X_{n+p}}(\sqrt{\alpha}) | X_n < q_{X_n}(\alpha)\right) \tag{53}$$

and $e^{x_0}(\alpha; k, k+p) = \lambda^{\sqrt{\cdot};x_0}(\alpha; k, k+p) - \sqrt{\alpha}$. We will limit ourselves to $k = 1$, since this would seem to be the most relevant case for day-to-day risk-management: one would want to know the after-effects of a possible value-at-risk violation tomorrow. Note that for large values of k it is to be expected that $e^{x_0}(\alpha; k, k+p) \simeq e(\alpha, p)$, at least for such values of a_1 for which the ARCH(1) has a unique (in distribution) stationary solution.

We estimated $e^{x_0}(\alpha; 1, 2)$ for a_1's between 0 and 2, with $x_0 = 1$ and $a_0 = 0.001$ and $\epsilon \sim t_4$ as before. It turns out that $\hat{e}^{x_0}(\alpha; 1, 2)$ now increases very rapidly for a_1's between 0 and 0.1, and then remains practically constant:

$a1$	$\hat{e}^{x_0=1}(0.05; 1, 2)$	std		$\hat{e}^{x_0=1}(0.01; 1, 2)$	std
0	0	0.02		0	0.03
0.01	0.06	0.02		0.12	0.04
0.02	0.11	0.02		0.2	0.05
0.03	0.14	0.02		0.23	0.05
0.04	0.16	0.02		0.25	0.05
0.05	0.17	0.02		0.26	0.05
0.06	0.18	0.02		0.28	0.04
0.07	0.18	0.02		0.27	0.04
0.08	0.19	0.02		0.28	0.05
0.09	0.19	0.02		0.28	0.05
0.1	0.19	0.02		0.28	0.05

Table 7.3 - Non-stationary excess tail probabilities as function of a_1.

For values of a_1 slightly bigger than 0.1, $\hat{e}(0.05; 1, 2)$ increases to 0.2 (0.02), and then stays this level. A similar thing happens with $\hat{e}(0.01; 1, 2)$, which becomes constant equal to 0.29 (0.05) just after $a_1 = 0.1$ - we checked this up to $a_1 = 10$, an unrealistically large value from the point of view of applications. Contrary to the stationary case, the conditional excess probabilities rise very quickly with a_1. However, neither $\hat{\pi}^{x_0}(0.05; 1, 2) = \hat{e}^{x_0}(0.05; 1, 2) + \sqrt{0.05} \simeq \hat{e}^{x_0}(0.05; 1, 2) + 0.2236$

or $\hat{\pi}^{x_0}(0.01; 1, 2) = \hat{e}^{x_0}(0.01; 1, 2) + 0.1$ come as close to the $\alpha \to 0$-limit of 0.5 as in the stationary case, for the range of a_1's considered.

We next look at the dependence of $e^1(1, 1+p) := e^{x_0=1}(\alpha; 1, 1+p)$ on the lag p, for $\alpha = 0.01$ and $a_1 = 0.05, 0.1, 0.2$ and 0.5; as before, the number in brackets is the standard deviation of our Monte Carlo estimate.

$p \backslash \hat{e}^1(1, 1+p)$	$a_1 = 0.05$	$a_1 = 0.1$	$a_1 = 0.2$	$a_1 = 0.5$
1	0.27 (0.05)	0.28 (0.05)	0.29 (0.05)	0.29 (0.05)
2	0.1 (0.04)	0.17 (0.04)	0.21 (0.04)	0.22 (0.05)
3	0.01 (0.03)	0.07 (0.04)	0.14 (0.04)	0.18 (0.05)
4	0.01 (0.03)	0.02 (0.03)	0.07 (0.04)	0.15 (0.04)

Table 7.4 - Non-stationary excess tail probabilites as function of the lag.

The excess probabilities start at roughly the same level, but their decay with increasing p is slower the larger a_1 is (for $a_1 = 0.05$ they are already indistinguishable from 0, within the precision of our MC calculations, from $p = 3$ onwards).

We finally take a look at the dependence on x_0. The next table gives an example, again with $\alpha = 0.01$, $k = 1$ and lag $p = 1$, and for three values of a_1:

$x_0 \backslash \hat{e}^{x_0}(1, 2$	$a_1 = 0.05$	$a_1 = 0.1$	$a_1 = 0.15$
0	0.07 (0.04)	0.11 (0.04)	0.14 (0.04)
0.1	0.1 (0.04)	0.16 (0.04)	0.18 (0.05)
0.2	0.14 (0.04)	0.21 (0.05)	0.23 (0.05)
0.3	0.18 (0.04)	0.24 (0.05)	0.26 (0.05)
0.4	0.20 (0.05)	0.25 (0.05)	0.27 (0.05)
0.5	0.22 (0.05)	0.26 (0.05)	0.27 (0.05)
0.6	0.24 (0.05)	0.28 (0.05)	0.28 (0.05)
0.7	0.25 (0.05)	0.28 (0.05)	0.28 (0.05)
0.8	0.25 (0.05)	0.28 (0.05)	0.28 (0.05)
0.9	0.26 (0.05)	0.28 (0.05)	0.28 (0.05)
1	0.27 (0.05)	0.28 (0.05)	0.29 (0.05)

Table 7.5 - Non-stationary excess tail probabilities as function of x_0.

For small x_0 there is a notable effect on the size of the excess probability, with initially an roughly linear increase which for large values of x_0 flattens out.

Appendix A. **Asymptotic expansions Mellin transforms**

We recall some basic facts regarding the Mellin transform, and its relation with asymptotic expansions of integrals. If $u : [0, \infty) \to \mathbb{R}$ is a locally integrable function which is bounded near 0, and has polynomial decay $|u(x)| \leq C|x|^{-k}$, then its Mellin transform $u^{\#}(s)$, defined by

$$u^{\#}(s) = \int_0^\infty u(x) x^{-s-1} dx, \quad \text{Re } s < 0. \tag{54}$$

The integral is absolutely convergent, and defines a holomorphic function on $\{s = \sigma + i\eta \in \mathbb{C} : -k < \sigma < 0\}$, which is bounded on each sub-strip $\{-k + \varepsilon < \sigma < -\varepsilon\}$, $\varepsilon > 0$.

If $u^{\#}(s)$ is integrable on the line $(\sigma - i\infty, \sigma + i\infty)$, $-k < \sigma < 0$, then we can recuperate u from $u^{\#}$ using Mellin's inversion formula:

$$u(x) = \frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} u^{\#}(s) x^s ds. \tag{55}$$

Integrability and other decay-properties of $u^{\#}(s)$ will generally follow from smoothness properties of u, using integration by parts. A very useful property of the Mellin transform is that it turns convolution on the multiplicative group $\mathbb{R}_{>0}$ into a product: if

$$u * v(x) := \int_0^\infty u(y)v(y^{-1}x)\frac{dy}{y}, \tag{56}$$

then

$$(u * v)^{\#}(s) = u^{\#}(s)v^{\#}(s), \tag{57}$$

whenever both sides make sense.

The (classical) connection between the Mellin transform of u and the asymptotics of $u(x)$ for $x \to \infty$ is based on the following observation: suppose that, under the above hypothesis, $u^{\#}(s)$, $s = \sigma + i\eta$, extends to a meromorphic function on a slightly larger strip $\{-k' < \sigma < 0\}$ (where $k' > k$), and has a single simple pole in $s = -k \in \mathbb{R}_{<0}$. Then by shifting the integration path of (55) from Re $s = \sigma$ to Re $s = -k - \varepsilon$, $\varepsilon < k' - k$, and formally applying Cauchy's residue theorem, we obtain that

$$\begin{aligned} u(x) &= \mathrm{Res}_{s=-k}\left[u^{\#}(s)x^{-s}\right] + (2\pi i)^{-1}\int_{\mathrm{Re}\ (s)=-k-\varepsilon} u^{\#}(s)x^{-s}ds \\ &= cx^{-k} + O(x^{-k-\varepsilon}), \end{aligned}$$

with

$$c = \mathrm{Res}\ _{s=-k}\left[u^{\#}(s)\right] = \lim_{s\to -k}(s+k)u^{\#}(s).$$

This can easily be made rigorous under the condition that $u^{\#}(s)$ is integrable on the line Re $(s) = -k - \varepsilon$ and that $|u^{\#}(\sigma + i\eta)| \to 0$ as $\eta \to \pm\infty$, uniformly for $s[-k-\varepsilon, k+\varepsilon]$. Multiple poles can be handled similarly, but will introduce logaritmic terms: if $u^{\#}(s)$ has a pole of order p in $s = -k$, then

$$\begin{aligned} \mathrm{Res}_{s=-k}\left[u^{\#}(s)x^{-s}\right] &= \lim_{s\to -k}\frac{d^{p-1}}{ds^{p-1}}\left((s+k)^p u^{\#}(s)x^s\right) \\ &= \sum_{\nu=0}^{p-1} c_\nu (\log x)^\nu x^{-k}. \end{aligned} \tag{58}$$

The coefficients c_ν can be computed by noting that if

$$u^{\#}(s) = \frac{a_{-p}}{(s+k)^p} + \frac{a_{-(p-1)}}{(s+k)^{p-1}} + \cdots + \frac{a_{-1}}{s+k} + a_0 + a_1(s+k) + \cdots,$$

is the Laurent expansion of $u^{\#}(s)$ around $s = -k$, then the residue in (58) is the coefficient of $(s+k)^{-1}$ in the Laurent expansion of the product $u^{\#}(s)x^s = x^{-k}u^{\#}(s)e^{(s+k)\log x}$. This gives

$$c_\nu = \frac{a_{-(\nu+1)}}{\nu!}. \tag{59}$$

In case there is more than one pole, one simply adds the contributions the individual poles. than one pole, we simply add up the contribution of each of the poles to the asymptotics. In particular, if $u^{\#}(\sigma + i\eta)$ extends merorphically to $\{-k - N < \sigma < 0\}$, with poles We summarize this discussion in the following lemma:

Lemma A.1. *(i) Suppose that, for suitable p, $c_\nu \in \mathbb{C}$ and $\eta > 0$, we have*

$$u(x) = \sum_{\nu=1}^{p-1} c_\nu (\log x)^\nu x^{-k} + O(x^{-k-\eta}), \quad x \to \infty. \tag{60}$$

Then $u^{\#}$ is of the form

$$u^{\#}(s) = \sum_{j=1}^{p} a_{-j}(s+k)^{-j} + h(s), \tag{61}$$

with $h = h(s)$ holomorphic on $-k-\eta < Re\ s < 0$.

(ii) Conversely, suppose that $u^{\#}(s)$ is given by (61), with $t \to u^{\#}(\sigma + it)$ integrable, and $|u^{\#}(\sigma + it)| \to 0$ as $t \to \pm\infty$, uniformly for σ in compact subsets of $(-k-\eta, 0)$. Then we have that, for all $\varepsilon > 0$, $u(x)$ has the truncated asymptotic expansion (60), with η replaced by $\eta - \varepsilon$.

Proof. (i) This follows from

$$u^{\#}(s) = \sum_{\nu=0}^{p-1} \int_1^{\infty} c_\nu \frac{(\log x)^\nu}{x^{k+s+1}} dx + h(s),$$

with $h(s)$ holomorphic on $-k-\alpha < \text{Re } s < 0$, and

$$\int_1^{\infty} \frac{(\log x)^\nu}{x^{\nu+1}} x^{-s} dx = (-1)^\nu \left(\frac{d}{ds}\right)^\nu \int_1^{\infty} \frac{dx}{x^{k+s+1}} = \frac{\nu!}{s+k}.$$

(ii) By shifting the integration path, as explained above. QED

References

[1] Acerbi, C., Tasche, D., 2002. On the coherence of expected shortfall, J. of Banking and Finance, 26, 1487-1503.

[2] Artzner, P., Delbaen, F., Eber , J. - M., Heath, D., 1999. Coherent measures of risk, Math. Finance, 9, 203-228.

[3] Bank for International Settlements website http://www.bis.org/bcbs/

[4] Basrak, B., Davis, R. A., Mikosch, T.,2002. Regular variation of GARCH processes, Stoch. Process. Appl., 99, 95-115.

[5] Berkowitz, J., O'Brien, J., 2002. How accurate are value-at-risk models at commercial banks? J. Finance LVII, 1093-1111.

[6] Bingham, N. H., Goldie, G. M., Teugels, J. L.,1987. *Regular Variation*, Cambridge University Press, Cambridge.

[7] Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity, J. of Econometrics 31, 307 - 325.

[8] Bollerslev, T., Engle, R. F., Nelson, D. B., 1994. ARCH models, in: Engle, R. F., McFadden, D., eds., *Handbook of Econometrics, vol. 4*, North Holland, Amsterdam.

[9] Brummelhuis, R. , Guégan, D. , 2001. Multi-period conditional distribution functions for conditionally normal GARCH(1, 1)-models, Univ. of Reims preprint; to appear in J. Appl. Probability, 2005.

[10] Campbell, J. Y., Lo, A. W., MacKinlay, A. C. (1997). *The Econometrics of Financial Markets*, Princeton University Press, Princeton, New Jersey.

[11] Embrechts, P., Klüppelberg, C., Mikosch, T., 1997. *Modelling Extremal Events for Insurance and Finance*, Springer, Berlin.

[12] Embrechts, P., McNeil, A., Straumann, D.: Correlation and dependence in risk management: properties and pitfalls In: Risk Management: Value at Risk and Beyond, ed. M.A.H. Dempster, Cambridge University Press, Cambridge, pp. 176-223

[13] Engle, R. F., 1982. Autoregressive conditional heteroskedacity with estimates of the variance of United Kingdom inflation, Econometrica 50, 907 - 1007.

[14] Engle, R. F. (Editor) (2000). *ARCH - Selected Readings*, Advanced Texts in Econometrics, Oxford University Press.

[15] Frey, R., McNeil, A., 2000. Estimation of tail-related risk measures for heteroscedastic financial time series : an extreme value approach, J. Empirical Finance, 7, 271 - 300.

[16] Goldie, C. M. , 1991. Implicit renewal theory and tails of solutions of random equations, Ann. Appl. Prob. 1, 126 - 166.

[17] de Haan, L., Resnick, S. I. , Rootzén, H., de Vries, C. G., 1989. Extremal behavior of solutions to a stochastic difference equation with applications to ARCH processes, Stoch. Process. Appl., 32, 213 - 224.

[18] Kesten, H., 1973. Random difference equations and renewal theory for products of random variables, Acta Math. 131, 207-248.
[19] Mandelbrot, B. (1963). The variation of certain speculative prices, J. Business, 36, 394 - 419.
[20] Nelson, D. B., 1990. Stationarity and persistence in the GARCH(1, 1) model, Econometric Theory, 6, 318-334.
[21] Rockafellar, A. T., Uryasev, S. (2002). Conditional value-at-risk for general loss distributions, J. Banking Finance, 26, 1443-1471.

Birkbeck College, School of Economics, Mathematics and Statistics, Malet Street Bloomsbury, London WC1E 7HX, UK, r.brummelhuis @statistics.bbk.ac.uk